



University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): Edward M. Smith, Joseph I. Hoffman, Laura E. Green and William Amos

Article Title: Preliminary association of microsatellite heterozygosity with footrot in domestic sheep

Year of publication: 2011

Link to published article:

<http://dx.doi.org/10.1016/j.livsci.2011.10.009>

Publisher statement: "NOTICE: this is the author's version of a work that was accepted for publication in Livestock Science. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Livestock Science, DOI: 10.1016/j.livsci.2011.10.009"

Preliminary association of microsatellite heterozygosity with footrot in domestic sheep

Edward M. Smith^a, Joseph. I. Hoffman^b, Laura E. Green^a and William Amos^c

^a*School of Life Sciences, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, UK*

^b*Department of Animal Behaviour, University of Bielefeld, Postfach 100131, 33501 Bielefeld, Germany*

^c*Department of Zoology, University of Cambridge, Downing Street, Cambridge, CB2 3EJ, UK*

Author email addresses:

Edward M. Smith: edward.smith@warwick.ac.uk

Joseph. I. Hoffman: j_i_hoffman@hotmail.com

Laura E. Green: Laura.Green@warwick.ac.uk

William Amos: wa100@hermes.cam.ac.uk

Corresponding author:

Edward M. Smith

Tel.: +44 (0) 24 7657 5874

Fax: +44 (0) 24 7652 3568

Abstract

Genetic heterozygosity in wild, unmanaged animal populations is often associated with protection against infectious disease. However, little is known about the relationship between heterozygosity and disease susceptibility in domesticated livestock, where disease resistance has the potential to improve animal welfare and productivity. We have investigated whether susceptibility to footrot, an important cause of poor welfare and reduced productivity in sheep, is associated with heterozygosity at 14 candidate microsatellite loci. Heterozygosity at locus BMC5221 was associated with resistance to footrot ($P=0.0034$). This locus was selected based on a gene ontology classification of 'response to Gram-negative bacteria'. Sheep homozygous at BMC5221 were at increased risk of virulent footrot (OR=4.8, 95% CI=1.5-15.3), with a dose response relationship between homozygosity and disease severity. A highly significant homozygote deficit was observed in sheep without virulent footrot (observed=4, expected=21, $\chi^2=13.76$, $P=0.0002$) but not in sheep that had clinical disease, suggesting homozygotes were disproportionately likely to contract virulent footrot. Our results indicate that genetic heterozygosity might be important for healthy immune function in domesticated livestock. The use of gene ontology codes might prove a useful strategy to target selection of candidate markers in future studies.

Keywords: Heterozygosity-fitness correlation; footrot; candidate loci; GO codes.

Introduction

Over 95% of sheep flocks in the UK have sheep with footrot (Kaler and Green 2008) and farmers estimate that 8-10% of sheep in a flock are affected at any one time (Kaler and Green 2009). The disease presents with a range of severities, from mild interdigital inflammation (interdigital dermatitis; ID), to separation of the hoof horn from the foot (Beveridge 1941), and all severities are associated with lameness (Kaler et al. 2009). The causative agent, *Dichelobacter nodosus* is an obligate anaerobe that can survive off the host for 7 days (Whittington 1995). Such a survival interval should be sufficient to allow all sheep in an infected population to be exposed, yet the majority does not develop disease, suggesting there is host variation in susceptibility.

Resistance to footrot varies between populations and phenotypes, and where heritability has been estimated, the majority of values lie in the range 0.10-0.20 (Nieuwhof et al. 2008), indicating the potential for effective selection for resistance. Indeed targeted breeding of sheep has led to a reduction in susceptibility to footrot in Corriedale sheep in New Zealand (Skerman and Moorhouse 1987). However, footrot is not in performance recording schemes, and recording and reporting occurrence of disease is not something all sheep breeders are keen to do. The identification of markers for resistance/susceptibility to footrot would therefore be a useful addition to conventional breeding (Nieuwhof et al. 2008). One such marker has been developed, based on haplotypes around the *DQA2* locus of the major histocompatibility complex (MHC), an important component of the immune system (Ennen et al. 2009; Escayg et al. 1997). Unfortunately such markers often appear population-specific, and this test is not effective in the breeds tested in the UK.

One promising avenue for the development of genetic markers for footrot is suggested by a host of recent studies that report heterozygosity-fitness correlations (HFCs). Here, heterozygosity

measured at a small number of microsatellite markers often predicts measures of fitness, including disease susceptibility. However, most research on genetic heterozygosity and disease resistance has focused on natural populations of animals, with high genetic heterozygosity generally associated with a reduced risk of infectious disease (Luikart et al. 2008; Rijks et al. 2008). Thus, individuals that are relatively homozygous at a panel of presumed neutral microsatellite markers tend to be at greater risk of infection and disease, with higher parasite loads and greater mortality risk (Rijks et al. 2008). Examples include bacterial and nematode parasites in sea lions and tuberculosis in wild boar (Acevedo-Whitehouse et al. 2005; Acevedo-Whitehouse et al. 2006). Although originally the underlying mechanism was thought to operate through inbreeding depression, recent theory and empirical studies suggest that in most cases the HFCs result through chance linkage between one or more of the markers and gene(s) with dominant or overdominant gene action (Balloux et al. 2004; Slate et al. 2004).

Unfortunately, there has been little research to date on domesticated populations (Kashtanov et al. 2003) where herd management and artificial selection impact on genetic variability. Thus the aim of the present study was to investigate whether genetic heterozygosity influences susceptibility to footrot in sheep. Using a targeted approach, we selected genetic markers specifically on the basis of proximity to genes that might be expected to play a role in resistance to infection. A panel of 14 candidate loci were developed and genotyped in a cohort of mule ewes (i.e. F1 crossbred animals) from one flock that was under constant observation for two years as part of a clinical trial of the management of lameness (Hawker 2007).

Materials and Methods

Study population and sampling

A lowland commercial flock from Oxfordshire, England, was studied from May 2005 – December 2006. The flock consisted of crossbred animals, with no information available on the parental

population, or specific sires. Ewes were inspected twice each week, their locomotion observed (Kaler et al. 2008), and sheep lame with locomotion score ≥ 2 were diagnosed and the cause of lameness recorded. A sheep was defined as having virulent footrot when there was separation of the hoof horn from the sensitive tissue of the foot with a characteristic smell. A sheep was defined as having interdigital dermatitis (ID) when the skin between the claws was reddened and moist. In 2006, a blood sample was collected into EDTA tubes under Home Office license and stored at -20°C from all ewes that were present on the farm and had been in the study for both years.

There were 191 blood samples from crossbred mule ewes with phenotypic data available for two years selected for analysis in two stages. In the first stage, sheep with maximally contrasting disease presentations were screened at 14 loci. This represented half the study population and included approximately equal numbers of case and control sheep, whilst also including the most severely affected individuals (defined as those affected with footrot in both years). The remainder of cases were sheep affected in one year only. Control sheep were selected to match the proportions of the case sheep, so approximately the same number of 'extreme' sheep (no ID, no footrot) were selected as there were sheep with footrot in both years, with the remainder made up of sheep with ID but not footrot. The resultant sample set consisted of 48 sheep with virulent footrot (CASES; 16 with virulent footrot in both years and 32 with virulent footrot in one year) and 47 sheep with no virulent footrot (CONTROLS; 15 with no signs of ID or virulent footrot and 32 with ID but not virulent footrot). In the second stage, the remaining, less contrasting samples were genotyped at the locus identified as most significantly associated with virulent footrot in the first stage. This consisted of 44 case individuals (1 with virulent footrot in both years and 43 with virulent footrot in one year) and 52 control individuals (21 with no virulent footrot and no ID and 31 with ID but not virulent footrot).

Marker selection

Fourteen candidate microsatellite loci were selected. These included markers which correspond to regions in the human genome where homozygosity has been associated with increased risk of mortality to bacterial disease (BMS1248, MCM148, ILSTS102; (Lyons et al. 2009)); markers located in the ovine MHC (OLADRBPs (Crawford et al. 1995)) and markers associated with other livestock diseases (CP26: tuberculosis (Acevedo-Whitehouse et al. 2005); INRA111 and INRA131: bovine mastitis (Schulman et al. 2004); MCMA16 and MCMA53: ovine scrapie (Luhken et al. 2007); TCRG4: ovine parasitism (Luikart et al. 2008); SHP4: ovine eczema (Phua et al. 1999); BL4: IgA QTL (Davies et al. 2006)). Two additional markers (TGLA48 and BMC5221) were chosen by identifying markers in the same linkage group as genes with the gene ontology (GO; www.geneontology.org) code for 'defence response to Gram-negative bacterium'. The locations of all non-ovine markers were aligned to the bovine genome using the comparative genomics function of the UCSC genome browser (<http://genome.ucsc.edu/>). Corresponding ovine loci were subsequently determined by alignment of the bovine and ovine genomes using the sheep comparative maps CMap site (<http://chimaera.vet.unimelb.edu.au/cmap/>).

Microsatellite genotyping

Genomic DNA was extracted from frozen blood samples using a DNeasy blood and tissue kit as recommended by the manufacturer (Qiagen, Crawley, UK). All loci were amplified using previously published and validated polymerase chain reaction (PCR) primers (see Supplementary data file 1 for details), as described in detail previously (Hoffman and Amos 2005). Briefly, PCR reactions were carried out in 12µL volumes containing 2µL template DNA (minimum of 10ng/µL), 1× Buffer A [10mM Tris-HCl (pH 8.3), 50mM KCl, 1.5mM MgCl₂, 0.01% Tween 20, 0.01% gelatin, 0.01% NP40], 1.0mM MgCl₂, 0.2mM each of dATP, dGTP and dTTP, 0.01mM dCTP, 2µM of each primer, 0.25U Taq polymerase and 0.01µCi [α^{32} P]-dCTP. Loci

were amplified using standard PCR cycling conditions of 94°C for 4min, 35 cycles of 45s at 94°C, 30s at T_1 , 30s at 72°C, and a final extension step of 72°C for 5min. PCR products were resolved by electrophoresis on standard 6% polyacrylamide sequencing gels, and detected by autoradiography. All gels were initially scored by one researcher (EMS) and checked for errors by a second investigator (JH). Primer sequences, genomic locations and annealing temperatures (T_1) of all markers are presented in Supplementary data file 1.

Statistical analyses

Observed and expected heterozygosity was calculated for each locus using 'Genepop on the Web' (Raymond and Rousset 1995). Deviations from Hardy-Weinberg equilibrium were assessed using Fisher's Exact Tests, with correction for multiple testing applied using the false discovery rate (FDR) implemented in Q-value (Storey and Tibshirani 2003a). Multilocus heterozygosity was quantified for each individual using the measure internal relatedness (IR; (Amos et al. 2001)), while for each marker it was scored as a binary measure (0 for homozygotes and 1 for heterozygotes). The association between heterozygosity and disease status in the first sample set (n=95) was investigated using generalised linear models (GLMs) in R2.7.2 (<http://www.r-project.org>). We initially tested whether IR was significantly associated with footrot disease status (positive (n=48) or negative (n=47)); then we tested the association at each locus individually. To compensate for overdispersion, significance testing was carried out using F-tests. Overdispersion is the norm for GLMs in which the response is modelled using the binomial. It is relatively simple to apply a conservative correction to take this into account by using F-tests rather than Chi-square tests. Resulting *P*-values were corrected for multiple tests using the FDR implemented in Q-value (Storey and Tibshirani 2003a).

After the second stage of genotyping, the association between heterozygosity and disease was calculated as described above in the full study population (n=175 successfully genotyped

sheep), because joint analysis of two-stage data is more efficient than replication analysis (Skol et al. 2006). We also investigated replication of the association in two cohorts containing equal proportions of sheep in each disease state (Table 1).

The risk of footrot for homozygotes at BMC5221 was determined by calculation of odds ratios (OR; (Bland and Altman 2000)), with analyses based on the full population dataset (n=175). The population was classified into four disease states: 1-virulent footrot in both years (n=15); 2-virulent footrot in one year (n=65); 3-ID but not virulent footrot (n=61); 4-no sign of ID or virulent footrot (n=34). The proportion of homozygotes at the associated locus in each of the four disease categories was calculated. The baseline for comparison was no sign of ID or virulent footrot. The overall OR was also calculated for the risk of footrot based on the binary disease classification (i.e. category 1+2 (n=80), compared with 3+4 (n=95)). The 95% OR confidence intervals were calculated as described previously (Bland and Altman 2000).

The observed and expected genotypic distribution at the associated locus was calculated using 'Genepop on the Web' (Raymond and Rousset 1995) with analyses based on the full population dataset (n=175). The probability that the distribution deviated from expectation was assessed using Chi-square (χ^2) tests and the CHIDIST function in Microsoft Excel.

Locus-specific signatures of selection were tested using an F_{ST} -outlier approach implemented in LOSITAN (Antao et al. 2008). F_{ST} -outlier tests are able to detect the majority of selected loci between populations in reasonably realistic conditions (Beaumont and Balding 2004). The initial population genotyped at all 14 loci was divided into two cohorts for analysis: virulent footrot positive (n=48) and virulent footrot negative (n=47). We ran 15,000 simulations as recommended (Antao et al. 2008) to achieve smoothing of confidence interval contours. Data were analysed using the infinite alleles model, a confidence interval of 0.95 and FDR of 0.1.

High F_{ST} at an individual locus compared to a simulation assuming neutrality would potentially indicate divergent selection whereas low F_{ST} suggests balancing selection.

Results

Association between homozygosity and footrot

The genotypes of all markers were in Hardy-Weinberg equilibrium. Observed heterozygosity ranged from 0.26 to 0.95 with an average of 6.6 alleles per locus (range 2-12; Table 2). Overall there was a non-significant negative association between heterozygosity, estimated as internal relatedness (IR) (Amos et al. 2001) and footrot disease status ($F_{[94]}=1.35$, $P=0.25$). However because we specifically targeted functional loci, further experimental effort is required to fully evaluate genome-wide effects of heterozygosity. We next tested each locus separately for individual association between heterozygosity and disease status. We found significant associations at two loci: TGLA48 ($F_{[91]}=4.42$, $P=0.036$) and BMC5221 ($F_{[90]}=8.35$, $P=0.0039$). Both of these loci were initially selected using gene ontology (GO) codes for response to Gram-negative infection (*D. nodosus* is Gram-negative). Following correction for multiple testing using the false discovery rate (FDR) (Storey and Tibshirani 2003a), BMC5221 remained significantly associated with virulent footrot ($Q=0.03$; Table 2). However, using the F_{ST} -outlier approach we were unable to detect any evidence of selection at the tested loci (see Supplementary data file 2), and all loci clustered around neutral F_{ST} .

Following the completion of the second round of genotyping, BMC5221 retained its strong association with virulent footrot ($F_{[174]}=8.59$, $P=0.0034$; $n=175$). Moreover, in the full sample set of 175 animals, homozygosity at BMC5221 decreased progressively from 26.7% in sheep with virulent footrot in both years of the study to 2.9% in sheep that had no clinical signs of virulent footrot or ID (Figure 1). The association was replicated when the population was analysed as

two cohorts matched for disease status (cohort 1: $F_{[87]}=4.36$, $P=0.037$; cohort 2: $F_{[86]}=4.23$, $P=0.040$; Table 1).

Risk of footrot based on genotype

Based on analysis of the full population dataset, compared to sheep never having clinical signs of footrot (n=34), sheep homozygous at BMC5221 (n=15) had an odds ratio (OR) of 12.0 (95% CI = 1.2-119.1; Figure 1) for having virulent footrot in both years. This fell to an OR of 6.0 (0.7-49.0, n=65) for footrot in one year, and 1.7 (0.2-17.1, n=61) for having ID but not virulent footrot. Overall the OR for having virulent footrot if homozygous at BMC5221 was 4.8 (95% CI=1.5-15.3) compared with ID or no clinical signs of disease.

Genotype distribution

Notwithstanding the lack of data on specific sire genotypes, which may have had varying levels of influence on our study population; the crossbred sheep in this flock which did not have virulent footrot were significantly more likely to be heterozygous at BMC5221 than expected (73 expected, 90 observed, $\chi^2=3.94$, 1 d.f., $P=0.04$). The number of homozygotes for individual alleles was too small to draw strong conclusions about specific allele effects. However, relative to Hardy-Weinberg expectations, there was a large deficit of individuals homozygous for allele 9 in both cases and controls (17 expected, 3 observed, $\chi^2=11.53$, 1 d.f., $P=0.0007$; Table 3). This contributes to the highly significant homozygote deficit in control sheep (21 expected, 4 observed, $\chi^2=13.59$, $P=0.0002$; Table 3). Among the cases, there were almost exactly the predicted number of homozygotes (15 expected, 14 observed, $\chi^2=0.08$, $P=0.78$; Table 3). However the distribution of these homozygotes differed from expectation ($\chi^2=35.32$, 7 d.f., $P=9.7\times10^{-6}$), with an excess of rare alleles, and only one at the commonest allele, allele 9 (Table 3).

Discussion

The aim of the present study was to investigate whether genetic variation influences susceptibility to footrot in domestic sheep. We have tested for possible associations between footrot in a flock of sheep examined several times each week over two consecutive years, and microsatellite homozygosity. We studied a commercial crossbred sheep population rather than a research flock to give our work a broader appeal; however without further testing we cannot guarantee that the population dynamics of this flock are representative of the UK sheep population. As such these results are provisional, but we would encourage the analysis of BMC5221 in any future research into the genetic basis for footrot susceptibility.

We were unable to detect a correlation with multilocus heterozygosity, or evidence for selection at any of the tested loci. However two putative associations were detected with individual markers, both of which were initially chosen for their proximity to genes involved with the response to Gram-negative bacteria. Despite a modest sample set, one of these markers, BMC5221, appears robust both to correction for multiple tests and to the inclusion of a second round of genotyping of relatively mildly affected individuals. It is perhaps unsurprising that there was no evidence of selection in our population. F_{ST} measures the reduction in heterozygosity in a subpopulation due, for example, to genetic drift and ranges from zero (no subdivision) to one (extreme subdivision). Our study sheep, in effect, formed an F_1 population where there has been no opportunity for genetic drift to occur, and the two cohorts analysed were drawn from the same overall population.

The observed genotype distribution at BMC5221 does not fit with a simple model where lambs are born with genotypes in Hardy-Weinberg Equilibrium and homozygotes preferentially contract footrot. Instead, disease-free animals exhibit a strong heterozygote excess not present in case animals. The most likely explanation for this is that selection of distantly related rams (mule

ewes are cross-bred) for breeding artificially inflates heterozygosity above Hardy-Weinberg expectations. Consequently, a heavily used sire with a rare allele would have undue influence on subsequent populations. An alternative hypothesis may be that there is disproportionate mortality in homozygotes in early life, for example because of other infectious diseases. In practice these alternative hypotheses are difficult to distinguish without testing the impact of breeding practices on heterozygosity.

The locus that was significantly associated with virulent footrot following correction for multiple testing (BMC5221) has characteristics that one might expect of a locus linked to a gene associated with disease susceptibility. Linkage disequilibrium should lead to genotypic disequilibrium between cases and controls, although not necessarily a pure 'heterozygote is good' signal. This is what we find, with both cases and controls deviating from Hardy-Weinberg expectations, though in different directions. The controls have a deficit of homozygotes, while the cases have the expected total number of homozygotes but not at the commonest alleles expected under Hardy-Weinberg equilibrium.

The strongest deviation from Hardy-Weinberg expectation involved a large deficit of allele 9 homozygotes in both case and control sheep. This might be due to purifying selection, though the implied mortality would have to be very high. As two of the three allele 9 homozygotes are present in the disease-free group, this genotype appears not to be associated with virulent footrot, but may impact on other diseases. Arguably, and more likely, is that allele 9 is present at a much higher frequency in one of the parent breeds compared with the other. This would facilitate the high allele 9 frequency in our population, and the low frequency of homozygote individuals. Quite why homozygotes for allele 5 are absent from the disease-free group (Table 3) is unclear, but this may again reflect a breed-specific allele frequency or, speculatively, linkage to a deleterious recessive allele.

The association between heterozygosity and disease is supported by the observed dose response relationship (Figure 1). Stratifying the data according to disease severity reveals a pattern where the association with homozygosity is at its highest among animals with the most severe symptoms and declines as the symptoms become weaker. This is exactly the pattern expected if disease severity reflects a balance between both genetic and non-genetic risk factors such that the most severely affected animals have a higher probability of carrying susceptible genotypes compared with those showing intermediate symptoms. This may also suggest that susceptibility to footrot and interdigital dermatitis is linked.

BMC5221 was selected because it is linked to the region containing the chromogranin precursor A gene (*CHGA*) on ovine chromosome (OAR) 18 (Maddox et al. 2001). However the genomic separation between *CHGA* and BMC5221 is quite large and the marker is actually located in a mitochondrial oxodicarboxylate carrier (*SLC25A21*), 11.5Mb upstream of *CHGA* (Dalrymple et al. 2007). OAR 18 is approximately 120cM long (Maddox et al. 2001), or, in terms of physical size, 72.5Mb, giving an average of 1.65cM/Mb. This suggests the separation between BMC5221 and *CHGA* is approximately 19cM. Whilst linkage disequilibrium (LD) patterns are breed-specific, 25% of microsatellite markers on OAR 18, separated by 10-20cM are in significant LD in Australian cross-bred sheep (Merino × Border Leicester) (Meadows et al. 2008). However, the proportion of markers in significant LD increases as the genetic distance separating them decreases, with 61% of markers separated by up to 5cM in significant LD (Meadows et al. 2008). In view of this, we used the virtual sheep genome (v.2; <http://www.livestockgenomics.csiro.au/perl/gbrowse.cgi/vsheep2/>) and bovine genome annotation to identify any other biologically plausible candidate genes in the region (Dalrymple et al. 2007). We identified two potential positional and functional candidate genes of interest: *NFKBIA* (nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha)

which is 1.6Mb upstream of BMC5221; and *PSMA6* (proteasome subunit α type 6) located 1.7Mb upstream of BMC5221. *NFKBIA* inhibits *NF- κ B* (nuclear factor kappa-light-chain-enhancer of activated B cells), which plays a key role in regulating the immune response to infection (Silverman and Maniatis 2001). It is activated following Toll-like receptor 4 (*TLR4*) recognition of Gram-negative bacteria (Beutler 2000), and so potentially has a role in the control of infectious disease. *PSMA6* also influences *NF- κ B*, with increased *PSMA6* expression apparently enhancing inflammation through activation of *NF- κ B* (Ozaki et al. 2006). Future studies should aim to characterise polymorphisms within these genes in sheep, and investigate associations with footrot.

Our approach of choosing candidate markers for footrot in domestic sheep in the UK has successfully identified at least one, and possibly two, novel loci associated with this disease. A variety of criteria were used to select markers initially, with the use of GO codes apparently proving successful. The focus on heterozygosity differs from previous investigations of genetic factors influencing susceptibility to footrot, which tend to search for heritable differences in susceptibility (Nieuwhof et al. 2008). Although effective at identifying additive genetic variability, such methods may overlook loci where heterozygosity is important, because heterozygosity is largely non-heritable. In this sense, our study complements existing techniques and also uncovers novel regions. Moreover, heritability studies of footrot often use clinical signs such as interdigital inflammation with no under-running of the hoof horn and so include a combination of footrot and ID (Nieuwhof et al. 2008), whereas we considered these presentations separately and detected a dose effect. By focusing on each trait it is possible that a purer genetic signal might be detected.

Although heterozygotes do not breed true, a genotype test could be used either to guide targeted mating, or to inform extra scrutiny of homozygotes. Thus, depending on the cost of a

test and the economic benefits likely to accrue, it might be worthwhile testing sheep to ascertain whether they are lower risk (heterozygote) or higher risk (homozygote), with scrutiny and (possibly) prophylactic treatment being directed more towards the latter. In terms of breeding, heterozygotes would be allowed unrestricted mating, since they will produce 50% heterozygous offspring regardless of their mate's genotype, whereas homozygotes could be matched to complementary homozygotes.

The phenotypic data used to characterise sheep in this study were more detailed than are usually available for investigations into the heritability of, or genetic susceptibility to, footrot in sheep. This was because of a two-year longitudinal epidemiological study of lameness in sheep that would be both costly and time-consuming to repeat. The data result in a high level of confidence in our phenotypic assignment of disease state. By restricting our analyses to a single sheep population, we also minimised genetic stratification in our sample. While future studies should aim to investigate greater numbers both of sheep and markers, it will be difficult to obtain such detailed phenotypic data for a large number of sheep.

Conclusion

We have identified a novel locus and two putative functional candidate genes associated with footrot in domestic sheep. So far we have examined a relatively modest numbers of markers and sheep. However, the GO code system could potentially be used to identify many more candidate loci. Greater knowledge of genetic susceptibility to ID and footrot could potentially be used to understand the pathogenesis of the disease and to guide breeding programmes to produce sheep with a naturally lower risk of disease.

376 **Supplementary data**

377 Supplementary data file 1: Genomic location, primer sequences and annealing temperatures of
378 markers genotyped in this study.

379 Supplementary data file 2: Graphical output from LOSITAN.

380

381 **Acknowledgements**

382 This work was supported by a Wellcome Trust ViP Fellowship awarded to E.M.S. and a Natural
383 Environment Research Council British Antarctic Survey Strategic Alliance Fellowship awarded to
384 J.I.H. The authors thank students and technicians for sample collection assistance.

References

- Acevedo-Whitehouse, K., Vicente, J., Gortazar, C., Hofle, U., Fernandez-de-Mera, I.G. and Amos, W. 2005. Genetic resistance to bovine tuberculosis in the Iberian wild boar. *Molecular Ecology*; 14, 3209-3217.
- Acevedo-Whitehouse, K., Spraker, T.R., Lyons, E., Melin, S.R., Gulland, F., Delong, R.L. and Amos, W. 2006. Contrasting effects of heterozygosity on survival and hookworm resistance in California sea lion pups. *Molecular Ecology*; 15, 1973-1982.
- Amos, W., Wilmer, J.W., Fullard, K., Burg, T.M., Croxall, J.P., Bloch, D. and Coulson, T. 2001. The influence of parental relatedness on reproductive success. *Proceedings of the Royal Society Series B: Biological Sciences*; 268, 2021-2027.
- Antao, T., Lopes, A., Lopes, R.J., Beja-Pereira, A. and Luikart, G. 2008. LOSITAN: a workbench to detect molecular adaptation based on a F_{st} -outlier method. *BMC Bioinformatics*; 9, 323.
- Balloux, F., Amos, W. and Coulson, T. 2004. Does heterozygosity estimate inbreeding in real populations? *Molecular Ecology*; 13, 3021-3031.
- Beaumont, M.A. and Balding, D.J. 2004. Identifying adaptive genetic divergence among populations from genome scans. *Mol Ecol*; 13, 969-980.
- Beutler, B. 2000. Tlr4: central component of the sole mammalian LPS sensor. *Current Opinion in Immunology*; 12, 20-26.
- Beveridge, W.I.B. 1941. Foot-rot in sheep: A transmissible disease due to infection with *Fusiformis nodosus* (n. sp.). Studies on its cause, epidemiology and control. CSIRO Australia; Bulletin No. 140, 1-56.

406 Bland, J.M. and Altman, D.G. 2000. Statistics notes. The odds ratio. British Medical Journal;
407 320, 1468.

408 Crawford, A.M., Dodds, K.G., Ede, A.J., Pierson, C.A., Montgomery, G.W., Garmonsway, H.G.,
409 Beattie, A.E., Davies, K., Maddox, J.F., Kappes, S.W. and et al. 1995. An autosomal genetic
410 linkage map of the sheep genome. Genetics; 140, 703-724.

411 Dalrymple, B.P., Kirkness, E.F., Nefedov, M., McWilliam, S., Ratnakumar, A., Barris, W., Zhao,
412 S., Shetty, J., Maddox, J.F., O'Grady, M., Nicholas, F., Crawford, A.M., Smith, T., de Jong, P.J.,
413 McEwan, J., Oddy, V.H. and Cockett, N.E. 2007. Using comparative genomics to reorder the
414 human genome sequence into a virtual sheep genome. Genome Biology; 8, R152.

415 Davies, G., Stear, M.J., Benothman, M., Abuagob, O., Kerr, A., Mitchell, S. and Bishop, S.C.
416 2006. Quantitative trait loci associated with parasitic infection in Scottish blackface sheep.
417 Heredity; 96, 252-258.

418 Ennen, S., Hamann, H., Distl, O., Hickford, J., Zhou, H. and Ganter, M. 2009. A field trial to
419 control ovine footrot via vaccination and genetic markers. Small Ruminant Research; 86, 22-25.

420 Escayg, A.P., Hickford, J.G. and Bullock, D.W. 1997. Association between alleles of the ovine
421 major histocompatibility complex and resistance to footrot. Research in Veterinary Science; 63,
422 283-287.

423 Hawker, L. 2007. An intervention study to minimise footrot in sheep. PhD Thesis. Dept. of
424 Biological Sciences, University of Warwick.

425 Hoffman, J.I. and Amos, W. 2005. Microsatellite genotyping errors: detection approaches,
426 common sources and consequences for paternal exclusion. Molecular Ecology; 14, 599-612.

427 Kaler, J. and Green, L.E. 2008. Naming and recognition of six foot lesions of sheep using written
 428 and pictorial information: a study of 809 English sheep farmers. Preventive Veterinary Medicine;
 429 83, 52-64.

430 Kaler, J. and Green, L.E. 2009. Farmers' practices and factors associated with the prevalence of
 431 all lameness and lameness attributed to interdigital dermatitis and footrot in sheep flocks in
 432 England in 2004. Preventive Veterinary Medicine; 92, 52-59.

433 Kaler, J., Wassink, G.J. and Green, L.E. 2008. The inter- and intra-observer reliability of a
 434 locomotion scoring scale for sheep. The Veterinary Journal; 180, 189-194.

435 Kaler, J., Daniels, S.L.S., Wright, J.L. and Green, L.E. 2009. Randomised clinical trial of long
 436 acting oxytetracycline, foot trimming and flunixin meglumine on time to recovery in sheep with
 437 footrot. J Vet Intern Med; Online early.

438 Kashtanov, S.N., Lazebnyi, O.E. and Gracheva, S.V. 2003. Signs of fitness and allozyme
 439 heterozygosity in an artificial population of sable *Martes zibellina* L. Russian Journal of
 440 Genetics; 39, 1693-1697.

441 Luhken, G., Zieleniewicz, D., Brandt, H.R., Buschmann, A., Groschup, M.H. and Erhardt, G.
 442 2007. Microsatellites MCMA53 and MCMA16 on OAR15 are associated with susceptibility to
 443 atypical scrapie. Animal Genetics; 38, 88-89.

444 Luikart, G., Pilgrim, K., Visty, J., Ezenwa, V.O. and Schwartz, M.K. 2008. Candidate gene
 445 microsatellite variation is associated with parasitism in wild bighorn sheep. Biology Letters; 4,
 446 228-231.

447 Lyons, E.J., Amos, W., Berkley, J.A., Mwangi, I., Shafi, M., Williams, T.N., Newton, C.R., Peshu,
 448 N., Marsh, K., Scott, J.A. and Hill, A.V. 2009. Homozygosity and risk of childhood death due to
 449 invasive bacterial disease. *BMC Medical Genetics*; 10, 55.

450 Maddox, J.F., Davies, K.P., Crawford, A.M., Hulme, D.J., Vaiman, D., Cribiu, E.P., Freking, B.A.,
 451 Beh, K.J., Cockett, N.E., Kang, N., Riffkin, C.D., Drinkwater, R., Moore, S.S., Dodds, K.G.,
 452 Lumsden, J.M., van Stijn, T.C., Phua, S.H., Adelson, D.L., Burkin, H.R., Broom, J.E., Buitkamp,
 453 J., Cambridge, L., Cushwa, W.T., Gerard, E., Galloway, S.M., Harrison, B., Hawken, R.J.,
 454 Hiendleder, S., Henry, H.M., Medrano, J.F., Paterson, K.A., Schibler, L., Stone, R.T. and van
 455 Hest, B. 2001. An enhanced linkage map of the sheep genome comprising more than 1000 loci.
 456 *Genome Research*; 11, 1275-1289.

457 Meadows, J.R., Chan, E.K. and Kijas, J.W. 2008. Linkage disequilibrium compared between five
 458 populations of domestic sheep. *BMC Genetics*; 9, 61.

459 Nieuwhof, G.J., Conington, J., Bunger, L., Haresign, W. and Bishop, S.C. 2008. Genetic and
 460 phenotypic aspects of foot lesion scores in sheep of different ages and breeds. *Animal*; 2, 1289-
 461 1296.

462 Ozaki, K., Sato, H., Iida, A., Mizuno, H., Nakamura, T., Miyamoto, Y., Takahashi, A., Tsunoda,
 463 T., Ikegawa, S., Kamatani, N., Hori, M., Nakamura, Y. and Tanaka, T. 2006. A functional SNP in
 464 *PSMA6* confers risk of myocardial infarction in the Japanese population. *Nature Genetics*; 38,
 465 921-925.

466 Phua, S.H., Dodds, K.G., Morris, C.A., Paterson, K.A., McEwan, J.C., Garmonsway, H.G.,
 467 Towers, N.R. and Crawford, A.M. 1999. Catalase gene is associated with facial eczema disease
 468 resistance in sheep. *Animal Genetics*; 30, 286-295.

469 Raymond, M. and Rousset, F. 1995. GENEPOP (version 1.2): population genetics software for
 470 exact tests and ecumenicism. *Journal of Heredity*; 86, 248-249.

471 Rijks, J.M., Hoffman, J.I., Kuiken, T., Osterhaus, A.D. and Amos, W. 2008. Heterozygosity and
 472 lungworm burden in harbour seals (*Phoca vitulina*). *Heredity*; 100, 587-593.

473 Schulman, N.F., Viitala, S.M., de Koning, D.J., Virta, J., Maki-Tanila, A. and Vilkki, J.H. 2004.
 474 Quantitative trait loci for health traits in Finnish Ayrshire cattle. *Journal of Dairy Science*; 87, 443-
 475 449.

476 Silverman, N. and Maniatis, T. 2001. NF-kappaB signaling pathways in mammalian and insect
 477 innate immunity. *Genes and Development*; 15, 2321-2342.

478 Skerman, T.M. and Moorhouse, S.R. 1987. Broomfield Corriedales: a strain of sheep selectively
 479 bred for resistance to footrot. *New Zealand Veterinary Journal*; 35, 101-106.

480 Skol, A.D., Scott, L.J., Abecasis, G.R. and Boehnke, M. 2006. Joint analysis is more efficient
 481 than replication-based analysis for two-stage genome-wide association studies. *Nature*
 482 *Genetics*; 38, 209-213.

483 Slate, J., David, P., Dodds, K.G., Veenvliet, B.A., Glass, B.C., Broad, T.E. and McEwan, J.C.
 484 2004. Understanding the relationship between the inbreeding coefficient and multilocus
 485 heterozygosity: theoretical expectations and empirical data. *Heredity*; 93, 255-265.

486 Storey, J.D. and Tibshirani, R. 2003a. Statistical significance for genomewide studies.
 487 *Proceedings of the National Academy of Sciences of the United States of America*; 100, 9440-
 488 9445.

489 Storey, J.D. and Tibshirani, R. 2003b. Statistical significance for genomewide studies. *Proc Natl*
 490 *Acad Sci U S A*; 100, 9440-9445.

491 Whittington, R.J. 1995. Observations on the indirect transmission of virulent ovine footrot in
492 sheep yards and its spread in sheep on unimproved pasture. Australian Veterinary Journal; 72,
493 132-134.

494

495

496

Figure caption

Figure 1. Proportion of homozygous individuals (grey bars), and the odds ratio (crosses) of disease, at BMC5221, by disease level.

The numbers above each category are the total numbers of sheep in that category; error bars represent 95% confidence intervals; ID: interdigital dermatitis.

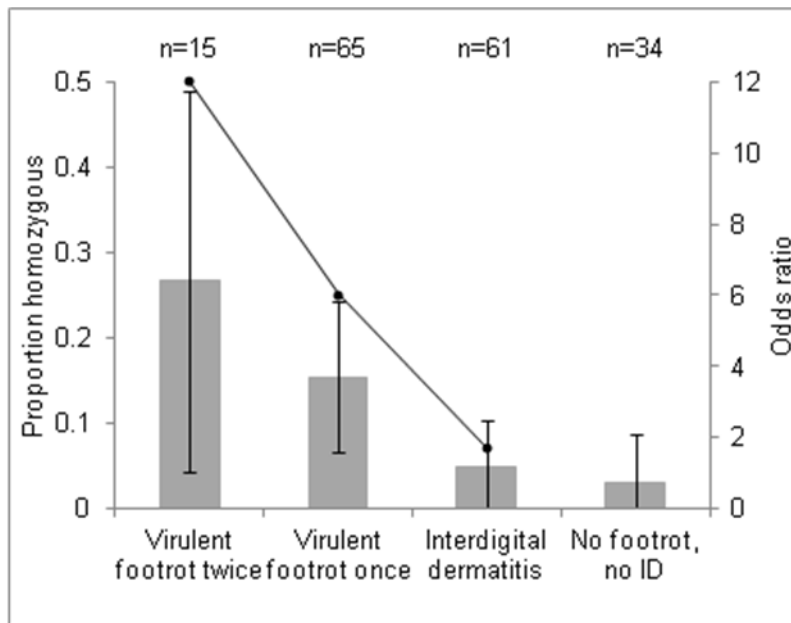


Table 1. Distribution of sheep in each disease state in the matched cohorts.

Disease status	Cohort 1	Cohort 2
Footrot in both years	7	8
Footrot in one year	33	32
ID, no footrot	31	30
No ID, no footrot	17	17

509

510 **Table 2.** Description of loci and results of GLMs of the effect of heterozygosity on footrot
511 disease status.

Name of	Reason for	Number of	Observed	<i>F</i>	<i>P</i>	<i>Q</i>
locus	selection	alleles	heterozygosity	value	value	value^a
SHP4	Ovine eczema	8	0.89	0.00	0.96	0.51
BMS1248	Human infection	9	0.70	0.00	0.95	0.51
TCRG4	Ovine parasitism	4	0.57	0.17	0.68	0.42
OLADRBPs	Ovine MHC	12	0.95	0.29	0.59	0.40
ILSTS102	Human infection	4	0.67	0.34	0.56	0.40
INRA131	Bovine mastitis	8	0.79	0.37	0.54	0.40
MCMA53	Atypical scrapie	7	0.71	0.39	0.53	0.40
MCMA16	Atypical scrapie	4	0.52	0.67	0.41	0.40
INRA111	Bovine mastitis	2	0.26	0.91	0.34	0.40
MCM148	Human infection	3	0.68	0.91	0.34	0.40
CP26	Porcine tuberculosis	6	0.78	0.92	0.34	0.40
BL4	IgA QTL	7	0.71	1.66	0.20	0.40
TGLA48	GO ontology	7	0.54	4.42	0.036	0.13
BMC5221	GO ontology	12	0.89	8.35	0.0039	0.03

512 ^aSignificance value adjusted for multiple testing using the false discovery rate implemented in Q-
513 value (Storey and Tibshirani 2003b).

514

515

516

517

518

519 **Table 3.** Differences between observed and expected numbers of homozygote individuals at
520 locus BMC5221 in sheep with and without virulent footrot.

Allele	-----Sheep with virulent footrot-----				----Sheep without virulent footrot----			
	Observed	Expected	χ^2	P value	Observed	Expected	χ^2	P value
1	5	3.08	1.20	0.27	1	2.49	0.89	0.35
2	0	0.00			0	0.00		
3	0	0.06	0.06	0.80	0	0.15	0.15	0.70
4	3	4.60	0.56	0.46	1	6.03	4.20	0.04
5	4	1.43	4.59	0.03	0	0.64	0.64	0.42
6	0	0.06	0.06	0.80	0	0.00		
7	0	0.22	0.22	0.64	0	0.35	0.35	0.55
8	1	0.04	24.85	6.2×10^{-7}	0	0.01	0.01	0.94
9	1	5.61	3.79	0.05	2	11.12	7.48	0.01
10	0	0.00			0	0.00		
11	0	0.00			0	0.03	0.03	0.86
12	0	0.00			0	0.00		
All	14	15.11	0.08	0.78	4	20.82	13.59	2.3×10^{-4}

521

522 Smith et al. Microsatellite heterozygosity and footrot.

523 **Supplementary data file 1:** Description of genotyped markers

524

Marker name	Forward primer	Reverse primer	Annealing temp. (T ₁)	Primer reference ¹	OAR ²	Ovine genomic location ³	
						OAR_start	OAR_end
SHP4	TATACCAGAATTTCTGGAGATAGG	CGGGATTACACTAGAATCAC	65	[1]	15		
BMS1248	GTAATGTAGCCTTTTGTGCCG	TCACCAACATGAGATAGTGTGC	58	[2]	3		
TCRG4	ATTTCTGGATTTGTAGAGCTCG	TGACAGTGTGCATGTGTGTTT	55	[3]	4		
OLADRBPs	CTGCCAATGCAGAGACACAAGA	GTCTGTCTCCTGTCTTGTGCATC	55	[4]	20		
ILSTS102	CAGGACTGAGTAACTAAGGC	AGGAGACAGCTACAAACCCC	55	[5]	24		
INRA131	GGTAAAATCCTGCAAAACACAG	TGACTGTATAGACTGAAGCAAC	65	[6]	3	75,329,813	75,330,023
MCMA53	CAACATGGATGGACCTATCTATGG	TCACAGAGTGGGACATGAATGA	65	[7]	15		
MCMA16	CTGCAAGGTATTCTGTTGTG	AGGAAAATGGCCAAATAGAG	65	[7]	15	2,322,013	2,322,559
INRA111	TTGTCGGTGTGGAGAAGCAC	GTTTCCCGGTAACCAATTCC	58	[6]	3	69,652,759	69,652,891
MCM148	ACACGGCTGAGTGACTGAACA	TATTCCTGCCCTGCAGATAGG	58	[8]	18		
CP26	GGCCTAACAGAATTCAGATGATGTTGC	GTCACCATACTGACGGCTGGTTCC	65	[9]	4	83,207,274	83,207,407
BL4	AAATTTTTCATCCTTCTTTCTGAC	TCACCCTGACTGTGAATGC	60	[8]	3	158,050,248	158,051,387
TGLA48	AAATGTTTTATCTTGACTACTAAGC	ACATGACTCTGCCATAGAGCAT	55	[10]	5	24,010,362	24,010,524
BMC5221	AGCAAGGAGAACAGGCATTC	CTTCTTTGGCAGCACAGTTTC	65	[11]	18	48,851,084	48,851,329

525

526 ¹May not be the first publication to describe the primer sequence

527 ²Sheep chromosome

528 ³Location data obtained from the virtual sheep genome v.2.0: <http://www.livestockgenomics.csiro.au/perl/gbrowse.cgi/vsheep2/>

529

530

531 References

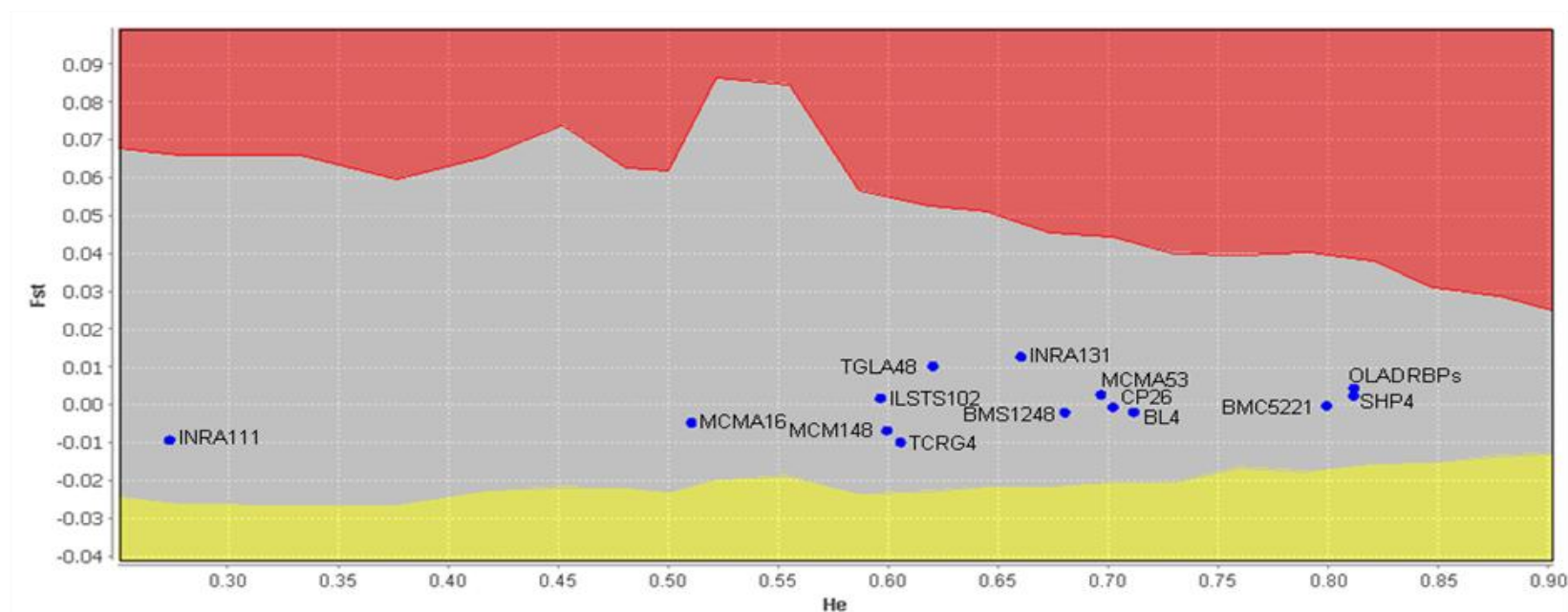
- 532 1. Phua SH, Dodds KG, Morris CA, Paterson KA, McEwan JC, Garmonsway HG, Towers NR, Crawford AM, 1999. Catalase
- 533 gene is associated with facial eczema disease resistance in sheep. *Anim Genet.* 30:286-95.
- 534 2. Stone RT, Pulido JC, Duyk GM, Kappes SM, Keele JW, Beattie CW, 1995. A small-insert bovine genomic library highly
- 535 enriched for microsatellite repeat sequences. *Mamm Genome.* 6:714-24.
- 536 3. ArkDB database: [<http://www.thearkdb.org/arkdb/>]

4. Crawford AM, Dodds KG, Ede AJ, Pierson CA, Montgomery GW, Garmonsway HG, Beattie AE, Davies K, Maddox JF, Kappes SW, et al. 1995. An autosomal genetic linkage map of the sheep genome. *Genetics*. 140:703-24.
5. Kemp SJ, Hishida O, Wambugu J, Rink A, Longeri ML, Ma RZ, Da Y, Lewin HA, Barendse W, Teale AJ, 1995. A panel of polymorphic bovine, ovine and caprine microsatellite markers. *Anim Genet*. 26:299-306.
6. Vaiman D, Mercier D, Moazami-Goudarzi K, Eggen A, Ciampolini R, Lépingle A, Velmala R, Kaukinen J, Varvio SL, Martin P, et al. 1994. A set of 99 cattle microsatellites: characterization, synteny mapping, and polymorphism. *Mamm Genome*. 5:288-97.
7. Australian sheep gene mapping website: [<http://rubens.its.unimelb.edu.au/~jillm/jill.htm>]
8. USDA MARC database: [<http://www.marc.usda.gov/genome/genome.html>]
9. Ede AJ, Pierson CA, Crawford AM, 1995. Ovine microsatellites at the OarCP9, OarCP16, OarCP20, OarCP21, OarCP23 and OarCP26 loci. *Anim Genet*. 26:129-30.
10. Ihara N, Takasuga A, Mizoshita K, Takeda H, Sugimoto M, Mizoguchi Y, Hirano T, Itoh T, Watanabe T, Reed KM, et al. 2004. A comprehensive genetic map of the cattle genome based on 3802 microsatellites. *Genome Res*. 14:1987-98.
11. Bishop MD, Kappes SM, Keele JW, Stone RT, Sunden SLF, Hawkins GA, Toldo SS, Fries R, Grosz MD, Yoo J, et al. 1994. A Genetic Linkage Map for Cattle. *Genetics* 136:619-639.

565
566

Smith et al. Microsatellite heterozygosity and footrot.

Supplementary data file 2: Graphical output from LOSITAN.



Fst = the measure of population subdivision ranging from 0 (no subdivision) – 1 (complete subdivision).

He = expected heterozygosity.

Blue dots represent the microsatellite loci investigated, and are labelled with their locus name. The red area is the region where loci potentially undergoing positive selection would be located; the grey area contains potentially neutral loci, and the yellow area would contain loci potentially subject to balancing selection.

567
568
569
570
571

572

573

574